

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
5 December 2002 (05.12.2002)

PCT

(10) International Publication Number
WO 02/097678 A1

(51) International Patent Classification⁷: **G06F 17/30**

(71) Applicant (for all designated States except US): **BODY1, INC.** [US/US]; 125 Cambridge Park Drive, Cambridge, MA 02140 (US).

(21) International Application Number: PCT/US02/09626

(22) International Filing Date: 29 March 2002 (29.03.2002)

(71) Applicants and

(72) Inventors: **MESSINA, Christopher, P.** [US/US]; 266 Western Avenue, Cambridge, MA 02139 (US). **SARHAN, Nabeil, O.** [US/US]; 276 Country View Drive, Warwick, RI 02886 (US). **KUGA, Shinpei** [US/US]; 70 Bishop Allen Drive #4, Cambridge, MA 02139 (US).

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/294,070 29 May 2001 (29.05.2001) US

(74) Agent: **VIKSINS, Ann, S.**; Schwegman, Lundberg, Woessner & Kluth, P.O. Box 2938, Minneapolis, MN 55402 (US).

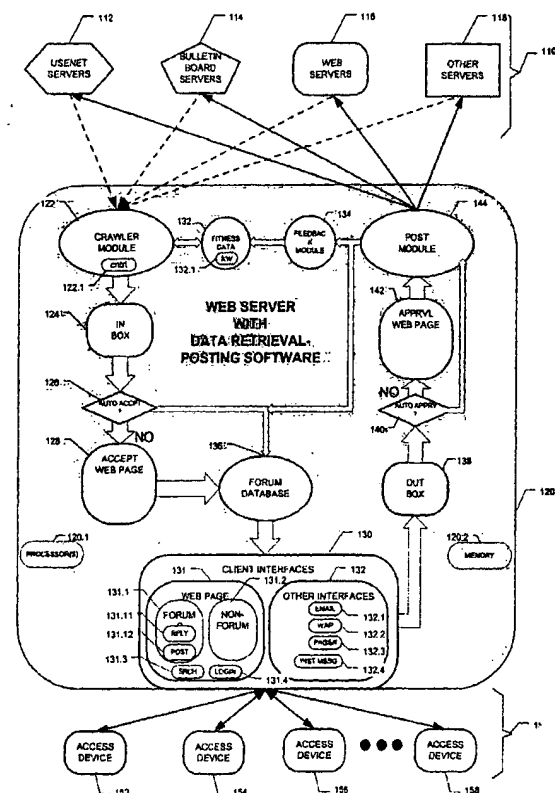
(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US PCT/US00/25886 (CIP)
Filed on 20 September 2000 (20.09.2000)
US Not furnished (CIP)
Filed on 20 March 2002 (20.03.2002)

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,

[Continued on next page]

(54) Title: METHODS, SYSTEMS, AND SOFTWARE FOR AUTOMATED GROWTH OF INTELLIGENT ON-LINE COMMUNITIES



(57) Abstract: In recent years, computers have become everyday communication tools, fast approaching the commonness of telephones and televisions. Driving this approach is rapid growth of the much-heralded Internet, which functions as a planetary communications system, enabling users to communicate with each other, to transmit data to each other, and to search for data of particular interest. Two problems stemming from the rapid Internet growth are the time and effort often necessary for users to find data of particular interest and the difficulty for users to find and communicate with other users who share interests in similar kinds of information. Accordingly, the present inventor devised systems, methods, and related software for encouraging and managing growth of databases, particularly theme-oriented databases, such as health-in-formation databases (Figure 5). One exemplary method entails establishing a theme-oriented database (520), receiving user contributions to the database, and granting users access rights to the database (152, 154, 156, 158) based on their contributions. The exemplary method also scores user contributions for quality and/or relevance to the theme of the database, and grants access rights based on the scores. Additionally, the exemplary method records the queries users of the database and facilitates communications between users having similar queries.

WO 02/097678 A1



MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,
SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,
VN, YU, ZA, ZM, ZW.

Published:

- with international search report
- with amended claims

- (84) **Designated States (regional):** ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Methods, Systems, and Software for Automated Growth of Intelligent On-Line Communities

Cross-Reference to Related Applications

5 The present application is a continuation of U.S. Provisional Application 60/294,070
which was filed on May 29, 2001, and a continuation-in-part of International Application No.
PCT/US00/25886 filed September 20, 1999 and published in English on March 29, 2001 as
Publication No. WO 01/22292. The international application is a continuation of U.S.
Provisional Application 60/154,885 filed on September 20, 1999. All these applications are
10 incorporated herein by reference.

Copyright Notice and Permission

One or more portions of this patent document may contain material subject to copyright
protection. The copyright owner has no objection to the facsimile reproduction by anyone of the
15 patent document, as it appears in the Patent and Trademark Office patent files or records, but
otherwise reserves all copyrights whatsoever. The following notice applies to this document:
Copyright © 1999-2002, Body1, Inc.

Background

20 Recent years have seen the explosive growth of the Internet as not only a communication
system of global scope, but also a global digital library. Indeed, the Internet has become a key
source of news and information on a full spectrum of topics for users the world over.

Fueling this growth, millions of users have submitted, or posted, information to services,
such as public newsgroups, message boards, and subscription mailing lists, that provide a
25 relatively simple and inexpensive mechanism for users to share information electronically with
many people. However, despite success in rapidly increasing the quantity of information on the
Internet, the present inventors have recognized that these services suffer from at least two

problems that ultimately limit the utility and value of their vast stores of user-generated information.

The first problem concerns the accessibility of the information stored by these services. Specifically, many services use their own unique form of user interface for posting and finding relevant information. This diversity, or lack of standardization, requires users of multiple services not only to learn different interfaces, but also to post and search messages on each service separately, even when the services focus on related topics. Although some interfaces allow the searching of several newsgroups for current and past articles, most, if not all, of them require users to repeatedly define and initiate searches to check for new relevant postings. Moreover, some services delete the oldest submissions or content on a weekly basis, reducing chances of users finding the content most relevant to their needs.

The second problem is the lack of an effective mechanism for collecting related information in a central database. In particular, many newsgroups and message boards focus on similar, highly related topics, yet they are not otherwise integrated or associated with each other or to other online sources of similar information, such as theme-oriented websites. For example, there are numerous independent health-oriented newsgroups, forums, message boards, and health-oriented websites in direct competition. This disunity forces users seeking the best user-generated information and potentially more authoritative, professionally authored information on a given topic to search multiple services and websites. Additionally, because many participants behave as spectators and do not actively contribute information, the direct competition not only leads to fragmented collection of information, but also divides the communities of active users across multiple services, frustrating the development of larger user communities.

Accordingly, there is a need not only to improve the value and utility of the vast stores of information on the Internet through new modes of access and collection, but also to encourage the growth of on-line communities.

Summary

To address this and other needs, the present inventors devised exemplary systems, methods, and software for automating the posting and retrieval of content from multiple data

services, such as newsgroups, message boards, forums, as well as encouraging active participation and growth of on-line communities.

For example, one system includes a posting module, a retrieval module, and a web server. The posting module, accessible through a webpage on the server, allows users to create and
5 initiate data postings that are sent automatically to several newsgroups, message boards, and/or other on-line information sources. The retrieval module automatically retrieves replies to the postings at each of the on-line sources and presents them through the webserver for user review and further reply, eliminating the need for users to repeatedly visit posting sites in search of reply messages. In addition to the retrieved replies, the retrieval module automatically finds and
10 retrieves content of any file type based on stored search or fitness criteria, ultimately enabling its integration into the content of one or more websites or databases for convenient user access.

Other aspects of this exemplary system include a feedback module for adapting fitness criteria over time to regulate the content of the databases according to user or administrative desires. Specifically, the exemplary system adapts the fitness criteria based on user interaction
15 with retrieved data and/or the content of user-generated postings, thus promoting evolution of website toward user interests. (specific user queries?)

Another system deploys the automated posting-and-retrieval capabilities within a webserver system that grants users data-access rights based on their contributions or submissions to one or more databases. This system scores the contributions based on quantity, quality,
20 and/or relevance, granting access rights based on the scores and thus promoting further contributions. Additionally, it records the queries of users of the database and facilitates communications between users having similar queries as well as users making similar contributions, further promoting growth of a communities of users around shared information.

25 **Brief Description of Drawings**

Figure 1 is a hybrid block-flow diagram of an exemplary data-retrieval-and-posting system 100 incorporating teachings of the present invention.

Figure 2 is a facsimile of a graphical user interface (or webpage) 200 suitable for use with system 100.

Figure 3 is a facsimile of a graphical user interface (or webpage) 300 suitable for use with system 100.

5 Figure 4 is a facsimile of a graphical user interface (or webpage) 400 suitable for use with system 100.

Figure 5 is a hybrid block-flow diagram of an exemplary system 500 based on system 100.

Figure 6 is a flow chart illustrating an exemplary method of operating system 100.

10 Figure 7 is a conceptual flow diagram illustrating various online discussion modes applicable to systems 100 and 500.

Detailed Description of Exemplary Embodiment(s)

The following description, which references and incorporates one or more figures, describes and illustrates one or more exemplary embodiments of the invention. These
15 embodiments, offered not to limit but only to exemplify and teach, are shown and described in sufficient detail to enable those skilled in the art to make and use the invention. Thus, where appropriate to avoid obscuring the invention, the description may omit certain information known to those of skill in the art.

20

Exemplary System Architecture and Method of Operation

Figure 1 shows an exemplary automatic data-posting-and-retrieval system 100. The connecting lines in the figure indicate that various components of the system are interconnected permanently or temporarily via a wired or wireless wide-area-network, such as the Internet, or a
25 secure wired or wireless local-area networks, such as a corporate intranet. They further indicate through use of arrowheads, whether communications over the connections are generally one-way or two-way. System 100 includes servers 110 (at top of figure), user access devices 150 (at bottom), and a web server system 120 (in the middle).

More particularly, servers 110 include one or more types of servers, such as one or more Usenet (NNTP or network-news-transfer protocol) servers 112, one or more bulletin or message-board servers 114, one or more web servers 116, and one or more other types of database servers 118, such as mail servers, messaging servers, and file servers. The invention, however, is not believed to be limited to any particular collection of servers or server types. Indeed, the invention is not limited to specific file type, since for example, various embodiments can retrieve graphics, video, audio, text, olfactory, and biofeedback files.

User access devices 150 include access devices 152, 154, 156, and 158 are representative. The term "access device," as used herein, encompasses browser-equipped personal computers, network appliances, personal digital assistants, telephones, cell phones, web phones, televisions, web television, etc. Thus, the present invention is not limited to any particular class or form of access device.

Selectively coupled between servers 110 and access devices 150 is webserver system 120 which not only automatically retrieves and post data from and to servers 110, but also provides a common interface structure between these servers and access devices 150.

Specifically, webserver system 120 includes, in addition to conventional processors 120.1 and memory 120.2 (shown in lower left and right corners), various software and data components which can take a variety of forms, such as coded instructions or data on an electrical, magnetic, optical, and/or magnetic carrier medium. These components, arranged to show an exemplary software architecture and methodology, include a crawler module 122 (upper left corner of box), inbox 124, decision block 126, acceptance webpage 128, client interfaces 130, fitness database 132, feedback module 134, and forum database 136. Additional components include outbox 138, decision block 140, approval page 142, and post module 144.

Crawler module 122 generally searches one or more external databases, accessible via servers 110, for data items or content based on search or fitness criteria defined in fitness database 132. The databases searched are defined in a control database 122.1.

The control database includes one or more site-, group-, or user-specific master subscription list of newsgroups and other servers, which are maintained by a system administrator, individual user, or automated software. Some embodiments can also automatically assess the fitness of new newsgroups and other network-accessible information sources based on a sampling of their content and then automatically add or recommend addition of newsgroups and sources that meet the given fitness or relevancy criteria to the control database or specific portion of the database. Newsgroup suggestions can also be triggered by looking at messages that others posted on multiple newsgroups, since these messages may suggest some similarity of newsgroup focus. (In addition, control database 122.1 stores control and state about a target information space. This is used to memorize state of the information space from the crawler module's point of view should the information space be dynamic or temporal.)

In the exemplary embodiment, crawler module 122 is implemented as a scheduled service (or daemon) that connects to a networked data servers via sockets, automatically sending appropriate access information (for example, usernames, passwords or other security credentials. Additionally, fitness database 132, which is initially supplied with at least one keyword or more generally content attribute 132.1 from an administrator and modified by feedback module 134, with each attribute associated with an indicator of whether its is desirable or undesirable, and a magnitude indicator of the relative strength or relative priority the crawler module should apply to the attribute in evaluating content for retrieval. The present invention is not limited to any particular type or class of content attributes. For example, some embodiments use attributes related to source code, object code, applets, particular binary data sequences, digital watermarks, file names, file sizes, file types, headers, tags, and metadata , markup languages, and semantic concepts.

More precisely, the crawler module uses a two-part screening process. The first evaluates the content prior to retrieval based solely on inclusion or exclusion of specific content attributes, such as keywords 132.1, defined in fitness database 132. (Some embodiments generate and retrieve summaries or facsimiles of found content rather than the found content in

its entirety.) The second filters the initially retrieved content data based on full-text keywords search using an objective fitness function having the form

$$TotalFitness = \sum_{i=1}^N weight(k_i) freq(k_i), \quad (1)$$

where N denotes the total number of keywords from the fitness database or relevant subportion of the fitness database that are in the text (or more generally content) being evaluated; $freq(k_i)$ denotes the frequency of the i -th keyword within the text being evaluated; and $weight(k_i)$ denotes the current weight assigned to the i -th keyword. This can be normalized or scaled any numerous ways, for example, by standard decimal scaling into a particular range, such as zero to one.

Some embodiments use non-linear fitness functions. For example, one embodiment uses the following fitness function:

$$TotalFitness' = \sum_{i=1}^N \left(weight(k_i) \sum_{j=1}^{freq(k_i)} j \right), \quad (2)$$

where the weight for each keyword is the same as in equation (1), but the effect of each keyword on $TotalFitness'$ increases non-linearly with each additional occurrence of the keyword. To illustrate, the following table compares the fitness contributions for various frequencies of a keyword that has an associated weight of 4, computed using equations (1) and (2):

Frequency	Linear Total Fitness	Non-Linear TotalFitness'
1	1(4)=4	1(4)=4
2	2(4)=8	(1+2)(4)=12
3	3(4)= 12	(1+2+3)(4)=24

In some embodiments, a configuration file or graphical interface feature allows administrators, users, or groups to select between use of equations (1) and (2) for fitness measurements. (Selection and adjustment of the fitness function, for example, via an

administrative interface, facilitates granular control over the search results.) Still other embodiments use other fitness functions and/or expand the range of selectable fitness-measurement options with other fitness functions based for example on combinations or variations of equations (1) and (2) and/or on weights that are functions of frequency and/or other document parameters. Some embodiments apply different fitness functions to different sources, source types, or file types.

Other embodiments associate one or more sets of keywords and/or weights in the fitness database with specific users or groups of users. (Users may be identified by their use of specific access credentials, such as username and password.) Additionally, some embodiments store user-defined Boolean and natural-language queries and/or necessary credentials and payment authorizations to access commercial or other restricted-access databases.

If the fitness for a particular document is greater than a threshold (a number provided during configuration, (most often set to 0, but can be made more restrictive for low memory environments) then the entire information file (text, source, time and other information associated with the content) is copied to inbox 124.

Inbox 124 simply aggregates all data retrieved from servers 110. However, some embodiments assign the data to one or more specific users or groups of users based on the user- or group-specific fitness criteria that led to its retrieval. Still other embodiments provide user- or group-specific inboxes.

Decision block 126 determines whether to automatically forward the contents of inbox 124 to forum database 136 or to queue it for further automated or manual screening. In the exemplary embodiment, this determination entails checking the status of an auto-acceptance flag. If the flag is true, the data is transferred to forum database 136, and if the flag is false, indicating no automatic acceptance, the data is incorporated into webpage 128 for manual screening.

Other embodiments use user-, group-, topic- or source-specific auto-acceptance flags. Still others combine user-, group-, topic- or source-specific auto-acceptance flags with temporal-window definitions to allow for automatic acceptance of certain defined content for certain users

during certain windows of time. And still other embodiments define the auto-acceptance flag as a function of not only user, group, topic, source, and time, but also additional relevance or legal criteria. (See discussion of auto-approval flag relative to decision block 140 below.)

Webpage 128 shows contents of inbox 124 and enables a website administrator or other
5 authorized entity, to individually accept or reject each retrieved item. This information can be viewed as a complete scrollable listing or a sublisting based on a fitness threshold or range of fitness scores. Webpage 128 also allows an administrator to alter fitness values of some information files, with any alterations ultimately translated to the keyword scores in fitness database 132.

10 Figure 2 shows an exemplary webpage or graphical user interface 200 used in an alternative embodiment of webpage 128. Interface 200 includes an inbox display region 210, a selected-content display region 220, administrative action region 230, and a fitness update region 230. Inbox display region 210 includes sort-command inputs 211-214, item markers 215, item date indicators 216, item identifiers 217, previous input, 218, and next input 219. Sort-
15 command inputs 211-214 invoke sorted display of inbox items by date, by source (or origin), by fitness score, and by thread. Item markers 215 are selectable to mark or designate one or more of the displayed items for further action. Item marker 215.1, representative of a selected marker, is filled in to denote its selected status. Item-date indicators 216 indicate respective dates, such as dates of retrieval or dates of original posting at a particular source, associated with the displayed
20 items. Item identifiers 217 which are selectable independent of item markers 216, present indicia of the content of the retrieved items current displayed in inbox display region 210. Item identifier 217.1, representative of a selected marker, is shaded to denotes its selection. Selected content display region 220 displays content and other data associated with selected data item in the inbox. Administrative action region 230 includes command inputs 231-237, which provide
25 options to accept all data items marked in the inbox or accept an entire thread associated with a selected data item.

Figure 3 shows another alternative webpage or graphical user interface 300 useful as an alternative or addition to webpage 128. Notably, webpage 300 includes an input 310 for

inputting a fitness level that causes display of all items having a fitness greater than the input fitness level, an input 320 for causing acceptance of a retrieved discussion thread into a specific designated category. Input 330 shows the computed fitness of the displayed item; however, this fitness may be changed by an administrator. (Such a change will ultimately result in
5 modification will be counted as a positive or negative event for relevant content attributes in the fitness database.) Region 340 shows markup language associated with or embedded within the displayed content.

As Figure 1 indicates, content that is accepted via an interface such as webpage 128, 200, or 300, is stored (or referenced via hyperlink or other pointer) in forum database 136 for
10 subsequent access via one or more clients (access devices) through client interfaces 130, which include a graphical-user interface, such as webpage 131 and other interfaces 132.

Webpage 131 presents the contents of forum database 136 in combination with other non-forum content to one or more specific users or groups of users. Specifically, webpage 131, which is associated with one or more theme-oriented or general-interest websites, includes a
15 forum region 131.1, a non-forum region 131.2, a search region 131.3, and a login region 131.4. Forum region 131.1 provides access to all or a user- or group-specific portion of forum database 136, and non-forum portion 131.2 provides access to other website content. (Figure 4 shows an exemplary webpage 400, which is displayed in response to forum region 130.1. and Figure 5 shows a webpage that is displayed in response to selection of an identifier for a retrieved item on
20 page 400)

Forum region 131.1 includes a reply region 131.11 and a post region 131.12. Selection of reply region 130.11 allows a user to reply to one or more messages in forum database without having to navigate to the source of the messages. And, selection of post region 131.12 allows a user to post original messages to one or more newsgroups, bulletin boards, and mailing lists.
25 In some embodiments, forum database 136 organizes the retrieved content into categories based on total fitness and/or particular keyword scores to assist users in locating specific content.

Search region 131.3 invokes a search interface on webpage 131 to allow users to search against already retrieved content or to command specific natural-language, Boolean, or

structured-query languages (SQL) searches against all the newsgroups or particular newsgroups or data sources using the crawler module or one or more separate search engine. In some embodiments, the search interface provides users the option to define target data and target date for receiving search results, such as overnight, next two days, or within a week, or within a month. During the specified time, the crawler module (or one or more other search engines) will search multiple sources across the internet for the target data. For some searches, the system will report search results via email, instant message, page, etc. Still other embodiments allow users to navigate through the forum database, select one or more particular items, and then invoke display of other retrieved items that are most similar to the selected items or command a search based on the contents of the selected items.

In addition to webpage 131, client interfaces includes other interfaces 132. This set of interfaces includes any other conventional or emerging interface structure or protocol recognition and conversion facilities necessary to communicate with any available or emerging form of access device. Thus, for example, other interfaces 132 includes an email interface 132.1 for communicating via email with one or more accesses devices, a wireless-access-protocol (WAP) 132.2 for communicating webpages to mobile devices, such as cell phones and personal-digital assistants; a pager interface 132.3 for communicating with pagers, and an instant-messaging interface for communicating using one or more instant-message protocols.

Outbox 138 serves as a first-in-first-out queue for outgoing reply and post messages from all users of the website(s) originating from client interfaces 130. However, other embodiments provide user- or group-specific outboxes. Outbox 138 outputs to decision block 140.

Decision block 140 determines whether to automatically forward the contents of outbox 138 to post module 144 or to queue it for further automated or manual screening. In the exemplary embodiment, this determination entails checking the status of an auto-approval flag, a preset true-false flag. However, other embodiments employ a more complex auto-approval process which entails measuring the fitness of the user original or supply post in a similar manner as used to evaluate incoming data within crawler module 122 and automatically approving postings that satisfy a threshold fitness level or fitness range.

Some of these embodiments relax the fitness criteria for internally generated postings (submissions) relative to the criteria for importing externally generated postings. Still others define the auto-approval flag as a function of user, group, topic, source, time, legality, and/or relevance criteria, including fitness data 132. Thus, the present invention is not limited in terms
5 of the type of auto-approval decision process. In any event, if the flag is true, the data is transferred to post module 144, and if the flag is false, indicating no automatic acceptance, the data is incorporated into webpage 142 for further screening.

Webpage 142 provides access to contents of outbox 138 and enables a website administrator or other authorized entity to individually screen reply and post items. In the
10 exemplary embodiment, webpage 142 follows a similar layout as webpage 200 or 300, as shown in Figures 2 and 3 respectively. In some embodiments, webpage 142 includes one or more of the following additional control features, beyond what is shown in Figures 2 and 3: query-author feature for sending a email to the author (assuming extraction of email address from message or public address book) or for querying forum content (or broader internet for other content from
15 this author; an evict-author feature for deleting all content from a particular author from the forum database and preventing retrieval of his or her items; a reject -similar feature for rejecting all items that are very similar to a marked or selected item based on a conventional semantic analysis, a check-legality feature for comparing the item to a listing of pirated or likely-to-be pirated content.

20 If an item is disapproved for posting, a system message is generated and transmitted to the author of the message. For disapproved postings that meet or appear likely to meet fitness criteria of other affiliated forums, websites, or databases (based on a list and other criteria held by webserver 120), some embodiments provide messages that refer users, via a hyperlink, to another website with its embodiment of system 100. Disapproved items, in some embodiments,
25 are posted externally as desired, but shunted to a separate area of forum database reserved for non-conforming content. Approved items are forwarded to post module 144.

Post module 144 distributes items, based on explicit or implicit approval, to one or more of servers 110 and forum database 136. In distributing the items, post module 144 performs any

necessary data and/or protocol conversions as necessary to ensure acceptance at each of the destinations. (Post module 144 also includes one or more hyperlinks, designating webpage 131, an associated recruitment page, affiliate websites, forums, etc.)

5 The exemplary embodiment implements post module 144 as a service (daemon) and that operates according to a schedule defined during configuration. For example, post module 144 can be scheduled to post hourly, daily, etc. Other embodiments drive post module on an event basis, such as immediately on receipt of a new item for posting or after accumulating a predetermined number of new items. In some embodiments, posted items originating from webpage 131 are posted directly to forum database 136, bypassing outbox 138 and the
10 subsequent approval process.

Post module 144 also identifies keywords and/or concepts in the posted items and forwards this information to feedback module 134.

Feedback module 134 modifies or adapts fitness criteria to reflect greater or lesser desirability of content having specific attributes. In the exemplary embodiment, adaptation
15 occurs in response to user or administrative actions on content containing specific keywords. Specifically, the exemplary embodiment treats user responses that reflect approval or disapproval of crawler-retrieved content as triggers to modify fitness criteria. Modifying the fitness criteria entails increasing or decreasing specific keyword weights or attribute weights, adding or deleting certain keywords, or more generally changing a parameter in a given fitness function or changing
20 fitness functions. For example, a user reply to a crawler-retrieved message or click on a link of a crawler-retrieved webpage signal content approval, whereas deletion or inattention to retrieved items signal disapproval.

Other embodiments collect more direct user feedback through polling or surveying users about regarding how useful a particular retrieved document is and then adjusting the fitness
25 criteria accordingly. For instance, one embodiment includes "more like this one," "less like this one," or "no more like this one" buttons in the frame of a content-display page to receive direct user input regarding desirability of particular content. Another embodiment uses a multi-point rating system. Common examples of such systems include four, five, and ten points.

Administrative actions, such as changing the fitness score of a retrieved item or rejecting data in the inbox can also signal an opportunity for adjustment of fitness criteria.

In response to a detected opportunity for adjustment, the exemplary embodiment generates and sends an information file containing text of the relevant document and any associated user-created documents to the fitness module. In turn, the feedback module not only searches the associated information file for keywords that are present in fitness database 132, but also identifies new keywords or other semantic attributes for inclusion in the database. The weights for preexistent keywords are then increased or decreased by some marginal amount or percentage based on a desired sensitivity or control loop dynamic. Weights for new keywords are initially set at the marginal weight and allowed to evolve to other keyword weights in response to user interactions with content containing the keywords or other weighted attributes.

One embodiment adjusts the weights of keywords based on positive or negative interactions with retrieved items. Specifically, after a predetermined number of positive or negative interactions with a specific retrieved (or original) item, the feedback module adjusts weights of all relevant keywords (or other attributes, such as source, author, header tags, size) by respective predetermined amount for positive and negative interactions. For example, if the threshold number of interactions for triggering an adjustment is 100 interactions, the positive adjustment is +1.1, and the negative adjustment is -0.9, every 100 positive or negative interactions would cause a 10% increase or decrease in the relevant attribute weights and thus a corresponding impact on the net growth or expansion of the forum or website content.

In some embodiments, administrators may also skew the fitness functions or keyword (or databases in specific ways for specific periods of time to encourage or discourage specific types of content. Thus, for example, for a political website, an administrator (or administrative program incorporating semantic analysis) might assess whether the forum database had a balance of content on both sides of an issue, such as increasing taxes, and adjust the fitness function or keyword weighting to establish or maintain balance. Alternatively, if balance was not desired in all or a portion of the forum database, the fitness function or attribute weighting could be

adjusted, again by an administrator or administrative program, to achieve the desired mix of content along one or more semantic dimensions.

Other embodiments also update the fitness data or criteria to include new keywords, concepts, or attributes of original and/or reply postings. The feedback module has the ability to
5 modify keyword weights based on the utility (fitness, number of replies, acceptance rate of similar posts) of a given post. In addition, the feedback module could also suggest new keywords by identifying similarities between posts deemed to be useful.

Exemplary Applications

10 Exemplary system 100 and its various components have countless applications. For example, system 100 can be used as an intelligence gathering mechanism by corporations, non-profits, individuals, and government agencies. For examples, corporations interested in assessing their public images or product appeal can set of a website using the exemplary software or a suitable variant to collect relevant postings from multiple forums, message boards,
15 etc. Moreover, the corporations may post specific content, such as advertising or product testimonials, to incite a "public" response. Additionally, the exemplary software can be used to post classified advertising for goods and services to multiple forums, to collect survey data, to conduct a multi-bid-point auction.

System 100 can also treat emails from one or more associated email servers as postings
20 or content for retrieval. From the perspective of posting, these embodiments monitor content from all or specific outgoing or incoming emails and modify the fitness functions to reflect a desirability of the content in these emails. Additionally, some embodiments include a further option to find new, previously undesignated recipients for the email (or posting) from a continually updated list of website members or newsgroups, mail lists, etc. Notably, this email
25 feature can also be deployed within a group or organization to identify relevant content from emails, such as job skills found on resumes or detect transfer of sensitive information.

Further applications include: supplemental subscription-based search service for web portals, such as Google or AltaVista; aggregation and organization of data of multiple file types across all or a subset of nodes in a corporate intranet to facilitate use of employee-generated data, such as address books, expense reports, etc; finder for rare items, missing persons, experts, professionals; search across company-specific portals for relevant online posts, such as resumes, new customers, or technical competencies.

10 Exemplary System 100 with Community-Building Software

Figure 5 shows an enhanced version of system 100 in the form of an exemplary webserver system 500. In addition to the components of system 100, system 500 includes membership database 510, theme database 520, and community-building software 530.

Membership database 510 includes information regarding members or potential members (users) of system 500. The figure shows this information as representative records 512 and 514, which are substantially identical in structure. In the exemplary embodiment, record 512 includes member-identification data 512.1, member-profile data 512.2, member-contact or -linking data 512.3, and member-participation data 512.4. Member-identification data 512.1 includes data for identifying or authenticating the identity of a user. Member-profile data 512.2 includes data describing the professional biographies and credentials of the member. Member-contact data 512.3 includes data, such as one or more postal addresses, telephone numbers, e-mail addresses, or URLs for facilitating contact or communications with the associated user. And, member-system-participation data 512.4 includes quantitative and qualitative information regarding actual and permitted use of the system by each user. For instance, the exemplary embodiment maintains one or more access scores for each member, indicating levels of access to respective portions of theme-oriented database 520.

Theme-oriented database 520 includes theme data 522, site data 524, query data 526, and user data 526. Theme data 522 includes one or more keywords, terms, concepts, or website

addresses which define one or more aspects of the thematic or topical content of database 520.

(In some embodiments, fitness data 132 and theme data 522 comprise the same data.)

Exemplary themes or topics includes general healthcare and wellness information for humans or other animals, such as dogs, cats, or fish; specific healthcare information for various parts of the human body, such as joints (knees, hips, elbows, spine, etc.) or organs (heart, lungs, stomach, kidney, liver, eyes, ears, skin, etc.); specific medical conditions, such as allergies (food, plant, etc.), cancer, arthritis, obesity, mental illness; auto-immune deficiency (HIV). Other exemplary topics include technology breakthroughs, health-technology breakthroughs, children, cooking, sports, entertainment, celebrities, politics, law, restaurants, consumer products, motion pictures, videos, music recordings, corporations, government officials, criminal activity, schools, science, wines, beers, foods, professional service providers (lawyers, doctors, contractors, artisans, etc.) colleges, alumni of educational institutions, genealogy, gossip, or sex. One exemplary health-oriented database includes user-generated health content, medical-journal content, and an archive of health-oriented feature stories. Thus, the present invention is not limited to any particular theme or class of themes.

Site data 524 includes feature articles, journal articles and other content added to database 524 manually by its creators, sponsors, or other parties governing or maintaining the database or automatically by the system itself. Query data 526 includes a listing of one or more queries (or query summaries) made by registered users or members of the community, against the database, with each query associated with one or more portions of the membership data for its submitting member. User data 526 includes user contributions to the database, with each contribution logically associated with or appended to one or more portions of the membership data for its submitting member. Site data 524 includes forum database 136.1 and user data 526 includes a user-specific forum database 136.2.

In its exemplary operation, system 500 not only facilitates posting, retrieval, and searching of content across multiple on-line data services, but also promotes the intelligent growth and development of a theme-oriented database or website by awarding users data-access

rights based on their data contributions and new-member referrals, and promoting further interaction of system users.

Exemplary System Operation

More specifically, Figure 6, which shows an exemplary flowchart 600, illustrates an
5 exemplary method of operating system 500. Flow chart 600 includes blocks 602-638, which are
executed serially in the exemplary embodiment. However, other embodiments of the invention
may change the order of execution and/or execute two or more blocks in parallel using multiple
processors or a single processor organized as two or more virtual machines or subprocessors.
Moreover, still other embodiments implement the blocks as two or more specific interconnected
10 hardware modules with related control and data signals communicated between and through the
modules. Thus, the exemplary process flow is applicable to software, firmware, and hardware
implementations.

The exemplary method begins at block 602, with automatically or manually establishing
an initial version of theme-oriented database 520. To this end, the exemplary method executes
15 blocks 602.1-602.4.

Block 602.1 entails determining a theme or topic by storing one or more keywords,
semantic concepts, or other content attributes in fitness data 132 and/or theme data 642. The data
in the fitness data and theme data are initially the same data or at least related to the same topic.
However, some embodiments may use dissimilar data. The theme or fitness data, in some
20 embodiments, identifies one or more key experts or authorities in a particular field. This allows
for retrieving content authored by these experts or content that references them and thus
potentially determining by association others experts and so on. Once the theme or topic of the
database is determined, execution continues at block 602.2.

Block 602.2 entails identifying a set of one or more candidate members. Exemplary
25 candidate members include existing websites or portions of websites related to the theme and
persons or firms with expertise or indicated interest in the theme. The publishers of identified
publications or websites are then invited to register as members of the system. The exemplary
embodiments identifies them by conducting searches based on the content attributes and/or

conducting searches based on relevant journals, research citations, articles, posts, periodical subscription lists, or expressed interest in electronic newsletters, requests in online or offline forums or bulletin boards. Identification can also include polling existing experts, contributors, and/or community members.

5 After identification of a set of candidates, the exemplary method evaluates them in block 602.3. This entails sorting them into three classes: experts, contributors, and community to facilitate prioritization of recruiting efforts. Although some embodiments employ subjective manual sorting techniques, the exemplary embodiment applies an quantitative algorithm that assesses the degree of overlap of candidate recommendations using a "6 degrees of separation"
10 model" which assumes that more widely known expert, contributors, or community members should have greater relative stature (access rights) with the website. An alternative algorithm determines a ranking based on the research citations, number, and quality of research papers, and number and quality of posts, number of press citations regarding the theme or fitness data, recency of relevant posts, informational requests, publications, or press citations.

15 Some embodiments use crawler module 122 capability to parse metadata attached to retrieved content together with its fitness algorithms to conduct quantitative and/or statistical analysis of content. In these embodiments, crawler module 122 targets content from one or more candidate members using its metadata, and measures and stores the fitness of the data, ultimately enabling one to readily determine statistics, such as the average fitness of content posted by a
20 given author within one or more specific on-line data services. More generally, the method could be used to assess the average fitness of all content for a particular online data service or a content author. The availability of such statistics facilitates comparison or ranking of content providers or authors in any number of desirable ways. Indeed, the highest ranking set of providers may become targets of high-priority recruitment efforts, and/or included with the
25 content attributes driving crawler-module retrievals to ensure a steady input of high-quality content to the website.

 Additionally, some embodiments use this methodology to cull poor-performing on-line data services from the search space of the crawler module and to suggest addition of new ones.

Still other embodiments may cluster other types of metadata, such as copyright, publisher, geographic origin of data, date of origin, around measured fitness levels, enabling use of conventional database querying to calculate a wide range of statistics, for example, average fitness of data posted from a specific geographic region. (Other embodiments with appropriate semantic analyzers may be able to infer meta data from the content itself.)

After identification, execution then continues at block 602.4. This block entails recruiting or inviting one or more of the expert, contributor, or community candidates to join the website. In the exemplary embodiment, the initial recruiting efforts include manual or automated transmission of e-mail, letter, or telephonic solicitations to join. Incentives, such as data- access rights, but a financial stake in the website community, reciprocal web-traffic opportunities, publicity or other promotion, and priority access to events or new information (information access rights)-- or combinations thereof. Others include direct financial compensation based on the quality, quantity, and relevance of contribution provided by the expert, contributor, or community member.

In other embodiments, recruiting entails crawler module 122 posting messages encouraging participants of newgroups, message boards, etc., from which it retrieves content, to visit or request information about one or more of its affiliated websites. In some cases, the messages includes hyperlinks or URLs to the site and some incentive for visiting the website, such as an electronic coupon or gift certificate or bonus access rights to otherwise inaccessible content on the site. Some embodiments even reply directly to specific retrieved postings with such messages. An exemplary message might state that based on the content of your post, you would be interested in the one or more of the following websites, on-line forums, newsgroups, etc.

In an event after initialization in block 602, execution of the exemplary method proceeds to block 604. Block 604, one or more of the candidate members establish a communications link with the system through webserver 120. This entails each of the candidate members using an access station, such as access device (or station) 152, to invoke the URL to the system. For example, the user at access station 152 would invoke "www.domain-name.com" to connect her

computer system (or other network appliance) to webserver system 120. After establishing the link to webserver 120 execution proceeds to block 604.

Block 606 entails receiving registration information from the candidate member. The registration information includes member-identification data, member-profile data, member-
5 contact or -linking data, and member-system-participation data. Member-identification data includes data for identifying or authenticating the identity of a user, such as a username and password. Member-profile information includes professional biographical information, such as present employment, professional achievements, educational or other promotional type material indicating or suggesting the authority or credibility of the registering member in the topic.
10 Member-contact data includes data, such as one or more postal addresses, telephone numbers, e-mail addresses, or URLs for facilitating contact or communications with the associated user. Member-system-participation data includes an access score that governs the level of access that the associated member has over theme-oriented database 520. The exemplary system determines an initial access score based on whether the user was referred by an existing member,
15 or whether the user was given a special invitation based on his or her expertise in the theme. If an existing member referred the user, the access score for the existing member is increased upon registration of the new user. After all registration information has been received for a particular member, exemplary execution proceeds to block 608

In block 608, the system records the received registration information in membership
20 database 510. Although the exemplary embodiment maintains membership database 510 separate from theme-oriented database 520 for heightened security, some embodiments combine the databases. With recording of the registration information, the exemplary method advances to block 610.

Block 610 entails the new member logging into the system to access theme-oriented
25 database 520. Specifically, this entails the new (or an existing) member manually or automatically entering a username and password, for example at login region 131.4 (shown as part of webpage 131 in Figure 1) in client interfaces 130. (Existing members bypass blocks 602-608 to reach block 610.) The username and password are then verified against those in

membership database 510. Affirmative verification advances the exemplary method to block 612.

In block 612, the system presents the member a home page, such as webpage 131 in Figure 1, for theme-oriented database 520. From the home page, the member decides to query database 520 or to contribute data to database 520 as indicated by decision block 652.

A member decision to make a contribution to the database branches execution to block 616, which entails receiving a contribution from the member. In this exemplary embodiment, contributions include posting or responding to a post at one or more on-line database services, such as forum data 136.1 within site data 644, or one or more other affiliated on-line services using post module 144. It also includes referring or inviting one or more new members to the website. In some embodiments, contributions also include responses to one or more survey or polling features on the website, or even clicking through to affiliate or sponsor websites. Execution then continues to block 618.

In block 618, the system evaluates or scores the contribution based on its quantity, quality and/or relevance to the theme-oriented content of database 520. To evaluate data contributions, the exemplary embodiment uses the evaluation logic of crawler module 122 and the fitness data in theme database 542 and/or fitness database 132. Another embodiment converts the contribution to a natural-language query and executes this query against all or part of database 520. The natural-language searching algorithm produces quantitative measures of the relevance of the contribution. Other embodiments produce the relevance or quality measures using inverse-document-frequencies factors that favor rare terms and/or frequency factors which favor terms that in the document to be scored. In some embodiments, the contribution is summarized using specialized software, such as that described in U.S. Patent 5,708,825, entitled Automatic Summary Page Creation and Hyperlink Generation, which is incorporated herein by reference. Other embodiments score the contribution based on additional factors, including for example, length (number of words), number of citations to leading journals, inclusion of hyperlinks to predetermined cites (such as sponsors of the system) and/or grammar. Still other embodiments score the contributions manually using an editorial board of experts in the theme. And still other

embodiments factor in the credibility or stature of the author based on previous contributions to the website and/or

Execution then continues at decision block 620, where the system determines if the evaluation was good or not. That is, is the contribution of sufficient relevance and quality to be
5 added to database 520. If the contribution has a good evaluation, based for example on its score exceeding some threshold value, the system proceeds to block 622.

In block 622, the system adds the contribution to database 520. In the exemplary embodiment, this entails appending the username along with one or more portions of the member's member-profile information and/or member-contact information to the beginning
10 and/or end of the contribution. (Some embodiments use a side by side presentation approach for the contribution and its attribution.) Thus, other members accessing this information can identify its contributing member and assess credibility and authority of the contribution. Moreover, if a contributing member has elected to allow publication of its contact information, such as its e-mail address, future users of the information may establish communications with the contributor.
15 Publication of a hyperlink or URL associated with the contributor will offer opportunities for reciprocal web traffic from webserver 120 to a webserver associated with the contributing member, creating an incentive for further contributions to the theme-oriented database.

In block 624, after adding the contribution to database 520, the system updates the access score portion of the participation data for the member based on the score of the contribution. In
20 the exemplary embodiment, this entails adding the score of the contribution to the initial or previous access score in membership database 510. However, other embodiments adopt other forms of update.

In block 626, the system offers the contributing member an opportunity to connect with one or more other members who have made similar contributions to database 520. In the
25 exemplary embodiment, this entails defining a query based on the contribution and executing this query against the user data portion of database 520, more precisely user data 529. Based on the results of this query, the exemplary embodiment presents one or more usernames and related contact or biographical information for other members who have made similar contributions to

the database, thereby facilitating communications between the contributing member and other contributing members.

On the other hand, if the contribution had a poor evaluation based on the threshold, execution proceeds to block 628 which outputs a message to the member indicating that the contribution was rejected. Some embodiments may offer an explanation for the rejection, and others refer the member to an alternate system with an appropriate theme for receiving the information. Still others reroute the contribution to the alternate system, automatically enrolling the member in the alternate system by transferring a copy of the associated membership information. After rejecting the data, execution returns to decision block 614, where the member can choose to add new data or to submit a query for database 520.

From block 614, a decision to query database 520 branches execution to block 630, which entails receiving a query from the member. Although the scope of the invention encompasses queries of any number of forms, the exemplary embodiment accepts queries as a textual string with Boolean connectors or as a natural-language query. (Moreover, the exemplary embodiment permits the member to restrict the query to specific portions of theme-oriented database 520, such as to site data or user data.) Execution of the exemplary method then proceeds to block 632.

In block 632, the system executes the query against database 520. In the exemplary embodiment, this entails running the query against the entirety of database 520. However, some embodiments restrict the query to one or more portions of database 520.

In block 634, the system presents query results to webserver 120 for viewing by the member through an access station. The exemplary embodiment presents the query results based not only on the query and the contents of database 520, but also on the access score for the member. Thus, for example, a low access score may result in all or a portion of the results being withheld from the member. Some embodiments advise the member quantitatively regarding the withheld portion of the results. For example, the system presents a message that a certain percentage of the results is withheld. Other embodiments present citations for the portions being withheld to assist the member in determining the desirability of this withheld information. Variations of this approach may present the profile of the contributors of the withheld results.

In presenting the results to the requesting member, the exemplary system arranges or organizes the results based not only on relevance to the query presented but also on factors, such as the access rights of the respective contributors of data to the system. Thus, for example, data contributions from contributors that have accrued relatively high access rights are generally presented before data contributions from contributors with relatively lower access rights, assuming the contribution are of approximately equal relevance to the query. This presentation mechanism further encourages members, particularly those with related websites, to contribute content to the database.

Some embodiments present the results in a predetermined order based on the portions of the database that contain them. For example, one embodiment presents found data in the order of feature articles, user contributions, and journal articles, with the items in each category arranged based on relevance and/or access rights of respective contributors.

In block 636, the system stores the query and associated member-profile and/or member-contact information to the query portion of database 520. Once

stored in database 520, the query is searched like any other content within the database. When query results include one or more queries the queries are presented along with the usernames associated with the queries.

In block 638, after storing the query, the system offers the contributing member an opportunity to connect with one or more other members who have made similar queries of database 520. In the exemplary embodiment, this entails defining a query based on the query and executing this query against the query data portion of database 520, more precisely user data 526. Based on the results of this query, the exemplary embodiment presents one or more usernames and related contact or biographical information for other members who have made similar contributions to the database, thereby facilitating communications between the member and other members with similar queries. Of course, the member then has the option to contact one or more of the other members. Other embodiments also presents the member options to connect with members who have published information relating to the query or to allow other members with similar questions to contact her in the future.

Figure 7 shows a conceptual flow diagram 700 of an exemplary method of fostering or developing online communities that are not limited to being hosted at one place at a particular time. In the diagram, the term "hosted community" refers to a community of users united in interest and by the central location of their discussion, such as at a data forum or website, but not united in time, meaning users or participants can join and leave a discussion at any time. The term "spontaneous community" refers to a community of users united in interest and time, but not by any particular central location. With use of the data-posting-and-retrieval system (denoted crawler in the figure), virtual communities of people or organizations with similar interest can be built dynamically through an automated process that identifies users of common interests, such as members of a theme-oriented website, automatically schedules an on-line meeting time, and sends invitations for a

spontaneous community discussions and/or hosted communities at one or more theme-oriented websites. These hosted or spontaneous communities can bring together users from public and/or private networks, such as users of a corporate intranet, members of a particular organization, or users from a particular geographic region. These electronic, discussion-based communities can unite users sharing common interests, organization memberships, or geographical locations by bringing together contributors of valuable, relevant content.

An exemplary method procedure for creating a discussion community entails using a crawler module to identify interests based on content and keyword searches across multiple discussion groups, compile contact information for the interested users (mobile phone, email, instant messaging, etc.), schedule a spontaneous discussion and send invitations notifying users of the time and venue of the spontaneous discussion and/or a hosted discussion. This provides users the option of participating in spontaneous discussion, or contributing to a hosted discussion, with contributors of valuable information.

In this manner, the posting and retrieval system can serve as an agent for compiling interest profiles in addition to relevant content. These interest profiles can be used in combination with a scheduling module to dynamically organize spontaneous electronic discussion-based communities, that can exist over fixed or wireless media via a variety of different platforms. These electronic, discussion-based communities can unite users sharing common interests, organization memberships, or geographical locations by bringing together contributors of valuable, relevant content.

Conclusion

In furtherance of the art, the present inventors have presented exemplary systems, methods, and software automatically post and retrieve data from multiple

online data services, such as newsgroups, message boards, etc, as well as encouraging growth of online intelligence communities by rewarding data contributions with access rights.

The embodiments described above are intended only to illustrate and teach one or more ways of making and using the present invention, not to restrict its breadth or scope. The actual scope of the invention, which embraces all ways of practicing or implementing the teachings of the invention, is defined only by one or more of the following claims and their equivalents.

Claims

1. A computerized system for automatically posting and retrieving information to and from sites (databases) on a computer network, the system comprising:
 - means for posting one or more first documents stored at a first site to one or more second sites in the computer network;
 - means for automatically collecting at the first site one or more second documents posted at one or more of the second sites in reply to one or more of the first posted documents; and
 - means for accessing the replies at a second site associated with a poster of the on the one or more first sites, including means for excluding one or more of the reply documents based on user- or administrator-defined criteria.
2. A computerized system for performing an automated sweep of a data network such as Usenet, in a continuous and realtime fashion, the system comprising:
 - means for searching through an entire network by leveraging computational power and bandwidth to rapidly identify, sort, and retrieve content from a network as it arrives
 - means for expanding the search area to include new sources (such as newsgroups, directories, bulletin boards, servers, etc.) as these sources arise
 - means for performing continuous background searches for files with specific attributes
2. A computerized system for automatically posting and retrieving information to and from sites on a computer network, the system comprising:

- means for automatically searching and retrieving documents, based on user-defined keywords or parameters, from one or more first sites on the computer network on a periodic, scheduled, or event-driven basis, including:
 - means for using results of past searches to improve future search results;
- means for posting one or more of the retrieved documents to a second site, including:
 - means for allowing users of the second site to post documents to one or more of the first sites; and
- means, responsive to users of the second site posting documents to one or more of the first sites, for retrieving and collecting reply documents that are associated with the posted documents on the one or more first sites, including means for excluding one or more of the reply documents based on user- or administrator-defined criteria.

2. The computerized system of claim 1, wherein the means for posting one or more of the retrieved documents, includes:

- means, responsive to an adjustable threshold, for including or excluding retrieved documents from the second site.

3. The computerized system of claim 1, wherein the means for allowing users of the second site to post documents to one or more of the first sites, includes:

- means for associating one or more of the posted documents with one or more of the retrieved documents.

4. The computerized system of claim 1, further including means for allowing a user of the second site to define or select the keywords and parameters and/or the first sites.

5. The computerized system of claim 1, wherein one or more of the posted documents is sponsored by a commercial entity.
6. The computerized system of claim 1, wherein one or more of the posted documents concerns an offer to sell merchandise or other private property.
7. The computerized system of claim 1, wherein one or more of the retrieved or posted documents comprises a hyperlink.
8. The system of claims 1, 2, 3, 4, 5, 6, or 7 wherein one or more of the recited means is implemented as a set of machine-executable instructions carried on an electronic, optical, or magnetic medium.
9. A computerized method or automatically posting and retrieving information to and from sites on a computer network, the method comprising:
 - automatically searching and retrieving documents, based on user-defined keywords or parameters, from one or more first sites on the computer network on a periodic, scheduled, or event-driven basis, including:
 - using results of past searches to improve future search results;
 - posting one or more of the retrieved documents to a second site, including:
 - receiving commands from users of the second site to post documents to one or more of the first sites; and
 - in response to the posting documents to one or more of the first sites, retrieving and collecting reply documents that are associated with the posted documents on the one or more first sites, including excluding one or more of the reply documents based on user- or administrator-defined criteria.

10. A machine-readable medium comprising instructions for implementing the method of claim 9.
11. A website or server including means for causing execution of the method of claim 9.
12. A personal computer system or other access device including means for causing execution of the method of claim 9.
13. A server comprising:
 - means for receiving data submissions from a plurality of users;
 - means for measuring the quality, relevance, or quantity of the received data submissions relative to a specified criteria;
 - means for automatically retrieving data from a first network accessible database, based on a specified criteria and storing the retrieved data in a second database associated with the server; and
 - means for granting one or more of the users access rights to data in the second database based on the measured quality, relevance, or quantity of the received data submissions.
14. A computerized system for collaborative online production and editing of multimedia or video projects, the system comprising:
 - interface means for producing, editing, and assembly of video or multimedia presentations that can be available either locally or online;
 - means for multiple users or user group to assemble and edit presentations comprising multimedia, sound, video, or graphics via online editing tools;
 - means for multiple groups or users to collaborate on the organization and assembly of multimedia clips via an Internet connection.

1. A method comprising exchanging access rights to a database for a data contribution to the database.
2. The method of claim 1, wherein the access rights exchanged for the database are based on the quantity, quality, or relevance of the data contribution.
3. A system comprising:
a network-accessible database; and
means for exchanging respective access rights to the database for respective data contributions to the database.
4. A method comprising:
accepting one or more respective data contributions to a database from one or more respective users; and
granting the one or more respective users access rights to the database based on quality or relevance of the respective data contributions.
5. A system comprising:
means for accepting one or more respective data contributions to a database from one or more respective users; and
means for granting the one or more respective users access rights to the database based on quality or relevance of the respective data contributions.
6. A system comprising:

means for receiving one or more respective sets of information from one or more respective users, with each user having a respective username and respective database-access rights;
means for storing the one or more respective sets of information in a database;
means for assigning at least one respective score to each of the respective sets of information; and
means for updating the respective database-access index for at least one user based on the scored assigned to the respective set of information.

7. The system of claim 6, further comprising means for serving data from the database to at least one user, based on a query and the respective database-access rights of the one user.
8. A method comprising:
receiving one or more respective sets of information from one or more respective users, with each user having a respective username and respective database-access rights;
storing the one or more respective sets of information in a database;
assigning at least one respective score to each of the respective sets of information; and
updating the respective database-access index for at least one user based on the scored assigned to the respective set of information.
9. The method of claim 8, further comprising transmitting data from the database to at least one user, based on a query and the respective database-access rights of the one user.

10. A method comprising:
accepting two or more queries or contributions to a database from two or more respective users; and
facilitating relationships between the users based on similarities of the queries or contributions.
11. A system comprising:
means for accepting two or more queries or contributions to a database from two or more respective users; and
means for facilitating relationships between the users based on similarities of the queries or contributions.
12. A method comprising:
accepting one or more first queries or contributions to a database from one or more first users;
accepting a second query or contribution to the database from a second user;
identifying one or more of the first users based on similarities of the second query to one or more of the first queries or contributions; and
facilitating a relationship between the second user and the identified one or more first users.
13. The method of claim 12, wherein facilitating a relationship between the second user and the identified one or more first users comprises providing address information for one or more of the first users to the second user.

14. The method of claim 13 wherein providing address information for one or more of the first users to the second user comprises determining whether one or more of the first users has granted permission for providing address information.
15. A system comprising:
 - means for accepting one or more first queries or contributions to a database from one or more first users;
 - means for accepting a second query or contribution to the database from a second user; and
 - means for identifying one or more of the first users based on similarities of the second query to one or more of the first queries or contributions; and
 - means for facilitating a relationship between the second user and the identified one or more first users.
16. A method comprising:
 - receiving one or more queries for a first database, with each query having an associated username;
 - storing the at least a portion of each of the queries in a second database, with each stored portion associated with the associated username;
 - receiving a new query for the first database from a user computer;
 - searching the second database based on at least a portion of the new query to identify one or more usernames; and
 - communicating an offer to access the identified one or more usernames to the user computer.*

17. The methods or systems of claims 1-16, wherein the queries or contributions concern medical or health information.

18. The methods or systems of claims 1-16, wherein the queries or contributions concern knees, hearts, lungs, allergies, HIV, specific diseases, specific body parts, cancer, lost children, cooking, sports, entertainment, celebrities, politics, law, restaurants, consumer products, motion pictures, videos, music recordings, corporations, government officials, criminal activity, schools,
5 science, wines, beers, foods, professional service providers, colleges, alumni of educational institutions, genealogy, gossip, or sex.
19. A computer-readable medium comprising instructions for facilitating execution of the methods of claims 1, 2, 4, 8, 9, 10, 12, 13, 14, or 16.

10

AMENDED CLAIMS

[received by the International Bureau on 23 August 2002 (23.08.02);
original claims 1-19 replaced by amended claims 1-28 (7 pages)]

What is claimed is:

1. A computerized system for automatically posting and retrieving
information to and from databases on a computer network, the system
5 comprising:
 - means for posting one or more first documents stored at a first database
to one or more second databases in the computer network;
 - means for automatically collecting at the first database one or more
second documents posted at one or more of the second databases in reply
10 to one or more of the first posted documents; and
 - means for providing access at the first database to one or more of the
automatically collected second documents, including means for
excluding one or more of the second documents, based on user- or
administrator-defined criteria.
- 15 2. A computerized system comprising:
 - a first database accessible via a computer network;
 - means for incorporating data, based on user- or administrator-defined
criteria, from one or more other databases into the first database; and
 - 20 • means for updating the user- or administrator-defined search criteria
based on user interaction with data at the first database.
3. A computerized system for automatically posting and retrieving
information to and from sites on a computer network, the system comprising:
25
 - means, responsive to user-defined keywords or parameters, for
automatically searching and retrieving documents from one or more first
sites on the computer network on a periodic, scheduled, or event-driven
basis, including:

- means for using results of past searches to improve search results;
 - means for posting one or more of the retrieved documents to a second site;
 - means for allowing users of the second site to post documents to one or more of the first sites; and
 - means, responsive to users of the second site posting documents to one or more of the first sites, for automatically retrieving and collecting reply documents that are associated with the posted documents on the one or more first sites, including means for excluding one or more of the reply documents based on user- or administrator-defined criteria.
4. The computerized system of claim 3, wherein the means for posting one or more of the retrieved documents, includes:
- means, responsive to an adjustable threshold, for including or excluding documents from the second site.
5. The computerized system of claim 3, wherein the means for allowing users of the second site to post documents to one or more of the first sites, includes:
- means for associating one or more of the posted documents with one or more of the retrieved documents.
6. The computerized system of claim 3, further including means for allowing a user of the second site to define or select the keywords and parameters and/or the first sites.
7. The computerized system of claim 3, wherein one or more of the posted documents is sponsored by a commercial entity.

8. The computerized system of claim 3, wherein one or more of the posted documents concerns an offer to sell merchandise or other private property.
- 5 9. The computerized system of claim 3, wherein one or more of the retrieved or posted documents comprises a hyperlink.
10. The system of claims 1-9 wherein one or more of the recited means is implemented as a set of machine-executable instructions carried on an electronic,
10 optical, or magnetic medium.
11. A computerized method of automatically posting and retrieving information to and from sites on a computer network, the method comprising:
- automatically searching and retrieving documents, based on user-defined
15 keywords or parameters, from one or more first sites on the computer network on a periodic, scheduled, or event-driven basis, including:
 - using results of past searches to improve future search results;
 - posting one or more of the retrieved documents to a second site, including:
 - 20 • receiving commands from users of the second site to post documents to one or more of the first sites; and
 - in response to the posting documents to one or more of the first sites, retrieving and collecting reply documents that are associated with the posted documents on the one or more first sites, including excluding one or more of
25 the reply documents based on user- or administrator-defined criteria.
12. A machine-readable medium comprising instructions for implementing the method of claim 11.

13. A website or server including means for causing execution of the method of claim 11.
- 5 14. A personal computer system or other access device including means for causing execution of the method of claim 11.
15. A server comprising:
- means for receiving data submissions from a plurality of users;
 - 10 • means for measuring the quality, relevance, or quantity of the received data submissions relative to a specified criteria;
 - means for automatically retrieving data from a first network accessible database, based on a specified criteria and storing the retrieved data in a second network accessible database associated with the server; and
 - 15 • means for granting one or more of the users access rights to data in the second database based on the measured quality, relevance, or quantity of the received data submissions.
- 20 16. A computerized system for collaborative online production and editing of multimedia or video projects, the system comprising:
- interface means for producing, editing, and assembly of video or multimedia presentations that can be available either locally or online;
 - 25 • means for multiple users or user group to assemble and edit presentations comprising multimedia, sound, video, or graphics via online editing tools;

- means for multiple groups or users to collaborate on the organization and assembly of multimedia clips via an Internet connection.

17. A computerized method comprising:

- 5 • automatically posting one or more first documents stored at a first database in a computer network to one or more second databases in the computer network;
- automatically collecting at the first database one or more second documents posted at one or more of the second databases in reply to one
- 10 or more of the first posted documents; and
- automatically granting users access rights to documents stored on the first database in return for their data contributions to the first database.

18. The method of claim 18, wherein the access rights to documents stored

15 on the first database are granted based on the quantity, quality, or relevance of the data contributions.

19. A system comprising:

- a first database accessible via a computer network;
- 20 • means for granting users access rights to the first database based on their respective data contributions to the database;
- means for posting one or more first documents stored at the first database to one or more second databases accessible via the computer network; and
- 25 • means for automatically collecting at the first database one or more second documents posted at one or more of the second databases in reply to one or more of the first posted documents.

20. The system of claim 19, wherein the access rights are granted to users based on quantity, quality, or relevance of their respective data contributions.
21. The system of claim 19, further comprising
- 5 • means for excluding one or more of the second documents from the first database, based on user- or administrator-defined criteria.
22. The system of claim 19, further comprising means for serving data from the first database to at least a first user, based on a first query and access
- 10 rights associated with the one user.
23. The system of claim 19, further comprising:
- means for receiving first data contribution or a first query from a first user of the first database and receiving a second data contribution or a
- 15 second query from a second user of the first database; and
- means, responsive to similarities of the first data contribution or first query to the second data contribution or second query, for facilitating a relationship between the first and second users.
- 20 24. The system of claim 19, wherein the means for facilitating a relationship between the first and second users includes means for communicating address information for the first user with the second user.
- 25 25. The system of claim 24 wherein the means for communicating address information for the first user to the second user comprises means for determining whether the first user has granted permission for providing the address information.

26. A computer-readable medium comprising instructions for execution of the method of claims 11,17, or 18.

27. The systems, methods, or media of claims 1-26, wherein the data,
5 documents, queries, submissions, or contributions concern medical or health information.

28. The systems, methods, or media of claims 1-26, wherein the data,
documents, queries, submissions, or contributions concern knees, hearts, lungs,
10 allergies, HIV, specific diseases, specific body parts, cancer, lost children, cooking, sports, entertainment, celebrities, politics, law, restaurants, consumer products, motion pictures, videos, music recordings, corporations, government officials, criminal activity, schools, science, wines, beers, foods, professional service providers, colleges, alumni of educational institutions, genealogy, gossip,
15 or sex.

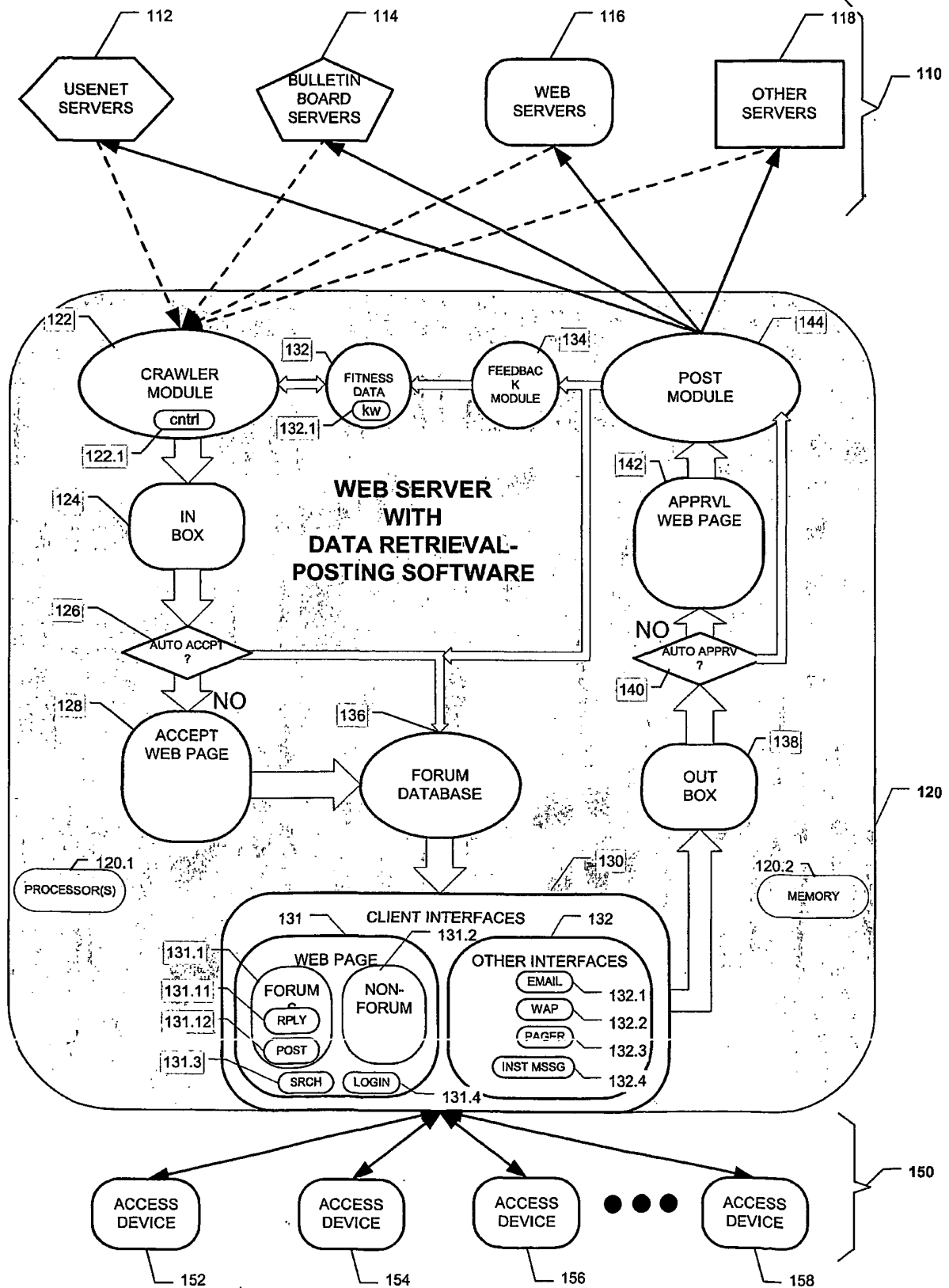


FIGURE 1

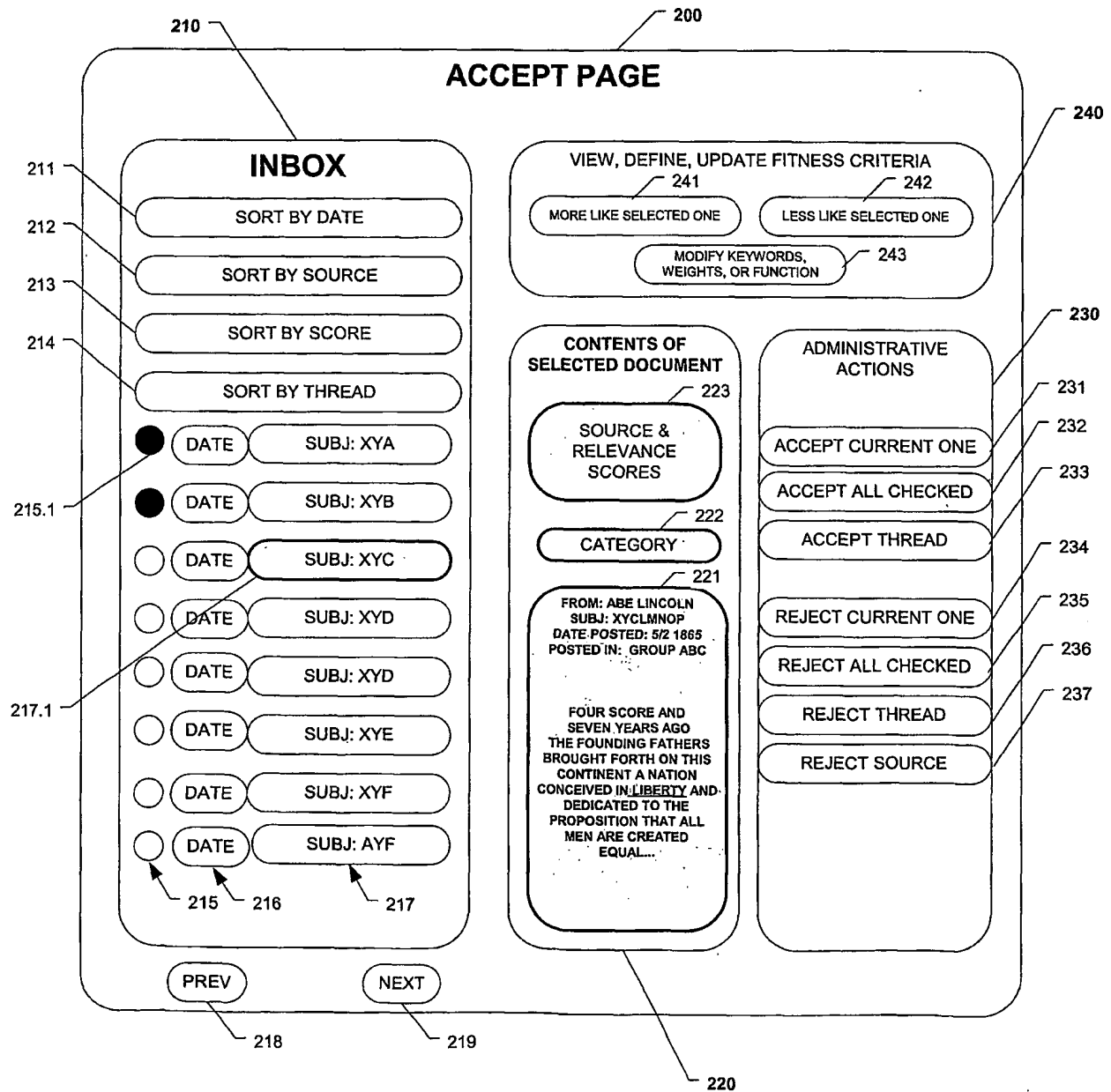


FIGURE 2

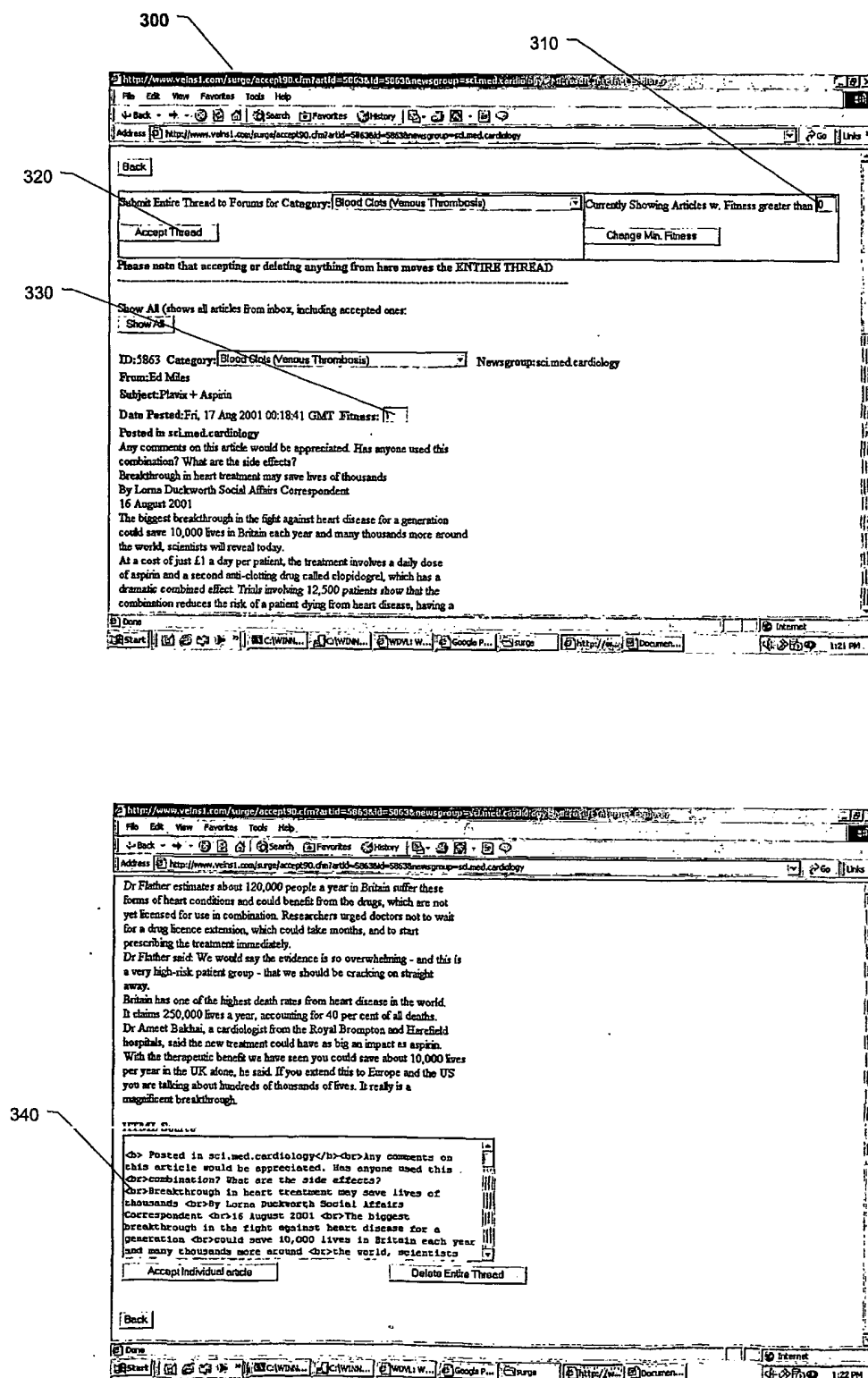


FIGURE 3

400

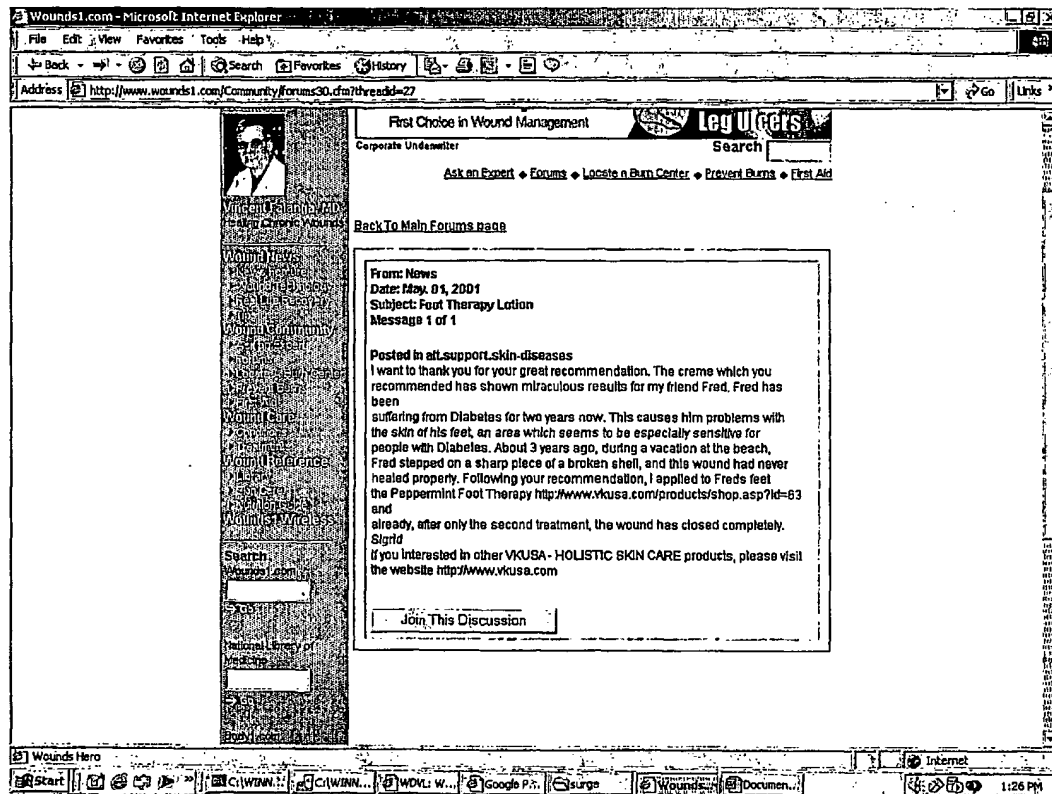


FIGURE 4

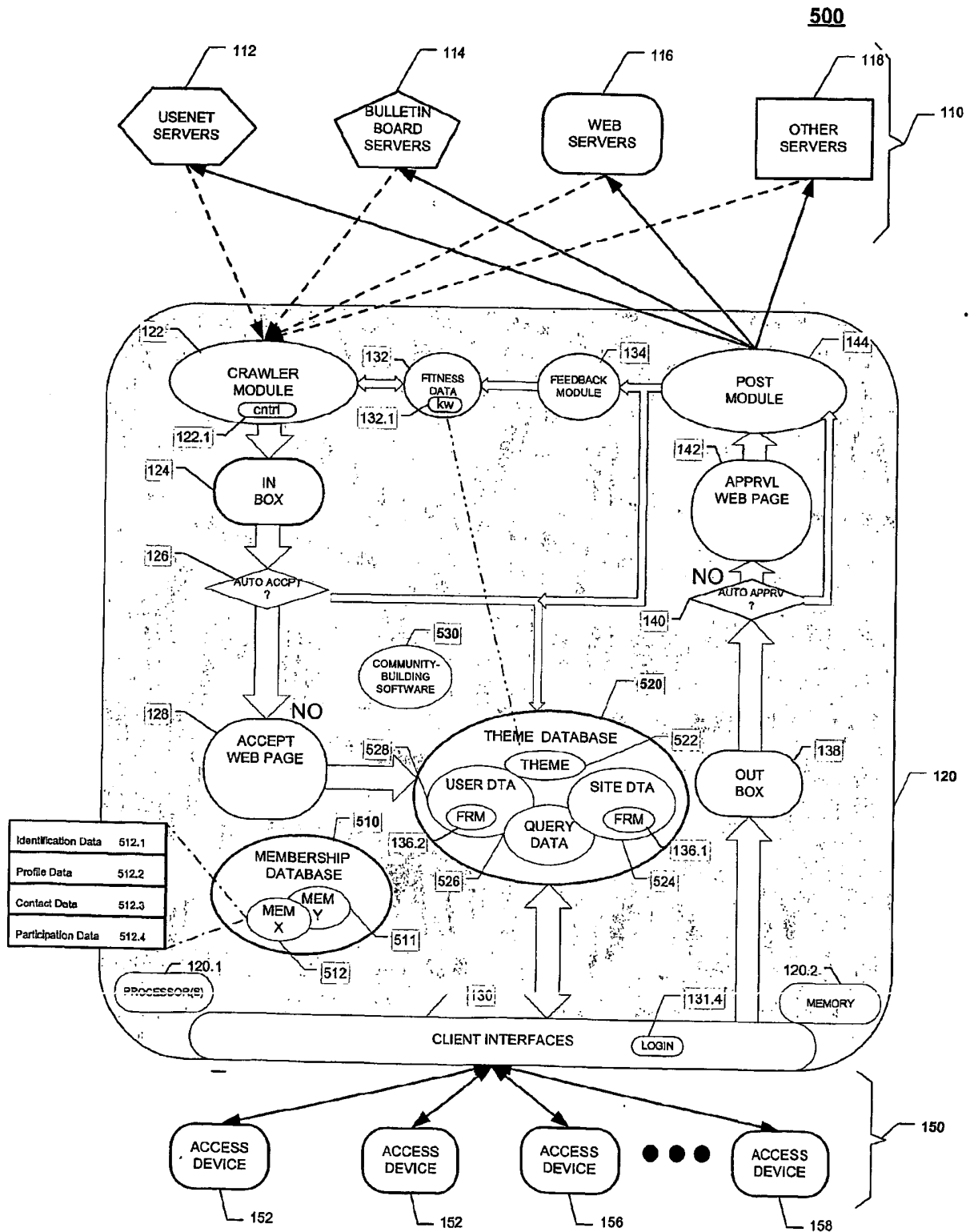
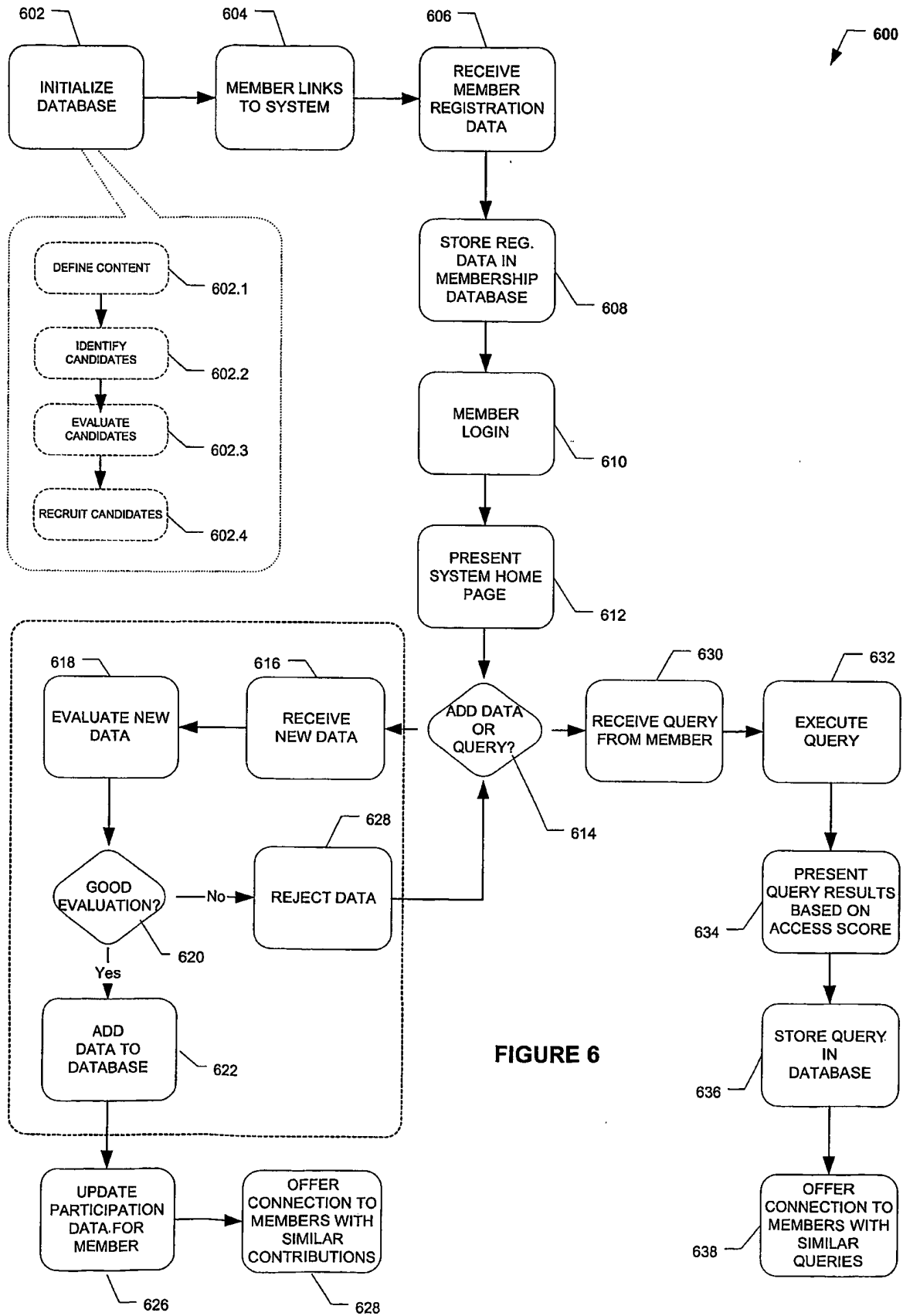


FIGURE 5



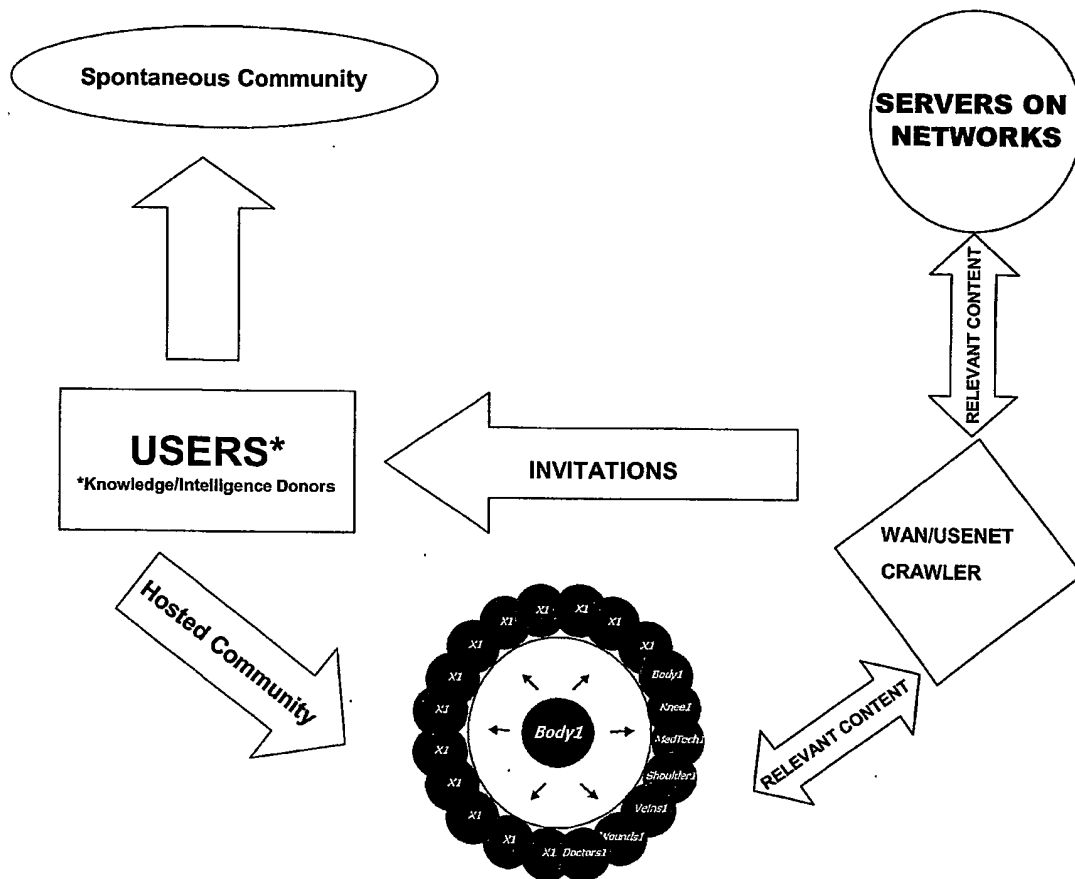


FIGURE 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/09626

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : G06F 17/30

US CL : 707/2, 3, 10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 707/2, 3, 10

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EAST

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 6,128,602 A (NORTHINGTON et al) 03 October 2000 (03.10.2000), abstract, column 2, line 29 - column 4, line 20.	1-2
Y	US 6,073,105 A (SUTCLIFFE et al) 06 June 2000 (06.06.2000), abstract, column 3, line 35 - column 8, line 40.	1-2

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent published on or after the international filing date	"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

20 May 2002 (20.05.2002)

Date of mailing of the international search report

28 JUN 2002

Name and mailing address of the ISA/US

Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703)305-3230

Authorized officer

Shahid Alam

Telephone No. (703)305-2358

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/09626

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claim Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
On page 30, claims are improperly numbered. Claim 2 marked twice. From bottom of page 30 to end of page 39, all of the claims are improperly marked.
2. ☐ Claim Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claim Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐
☐

The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/09626

Box III TEXT OF THE ABSTRACT (Continuation of Item 5 of the first sheet)

NEW ABSTRACT: In recent years, computers have become everyday communication tools, fast approaching the commonness of telephones and televisions. Driving this approach is rapid growth of the much-heralded Internet, which functions as a planetary communications system, enabling users to communicate with each other, to transmit data to each other, and to search for data of particular interest. Two problems stemming from the rapid Internet growth are the time and effort often necessary for users to find data of particular interest and the difficulty for users to find and communicate with other users who share interests in similar kinds of information. Accordingly, the present inventor devised systems, methods, and related software for encouraging and managing growth of databases, particularly theme-oriented databases, such as health-in-formation databases (Figure 5). One exemplary method entails establishing a theme-oriented database (520), receiving user contributions to the database, and granting users access rights to the database (152, 154, 156, 158) based on their contributions. The exemplary method also scores user contributions for quality and/or relevance to the theme of the database, and grants access rights based on the scores. Additionally, the exemplary method records the queries users of the database and facilitates communications between users having similar queries.

THIS PAGE BLANK (USPTO)